Physics 54

Wave Optics 1

Not only is the universe stranger than we think — it is stranger than we can think.

— Werner Heisenberg

Superposition of Harmonic Waves

The essential characteristic of energy transport by waves is that waves obey the **superposition principle**. This means that two waves in the same spatial region can *interfere*, rearranging the energy in space in a pattern often quite different from that of either wave alone. Since light propagates as a wave, we will analyze this phenomenon.

We begin with a mathematical problem: How do we find the wavefunction for the combined wave resulting from interference of two waves?

We consider two harmonic e-m waves of the same frequency and wavelength, both moving along the x-axis, but differing in phase by the angle δ. The wavefunctions are the expressions describing the E-fields. We will assume the fields to oscillate in the y-direction, so by E we mean the y-component of $\mathbf{E}$. We have for the two waves:

$$E_1(x,t) = E_{0_1}\cos(kx-\omega t)$$
$$E_2(x,t) = E_{0_2}\cos(kx-\omega t+\delta)$$

The superposition of these waves will result in another harmonic wave with the same frequency and wavelength, moving in the x-direction, so $E = E_1 + E_2$ must have the form

$$E(x,t) = E_0\cos(kx-\omega t+\phi)$$

Our problem is to find the constants E_0 and ϕ in terms of the amplitudes of the original waves and the phase difference δ. Our main interest is in the *intensity* of the resulting wave, which is proportional to ${E_0}^2$; we are usually less interested in ϕ.

To solve this problem we employ a trick based on a famous mathematical theorem:

Euler's theorem	$e^{i\theta} = \cos\theta + i\sin\theta$

Here $i = \sqrt{-1}$ is the imaginary unit. This remarkable formula says that exponential functions and trigonometric functions are related through complex numbers. Let us first review of some facts about complex numbers.

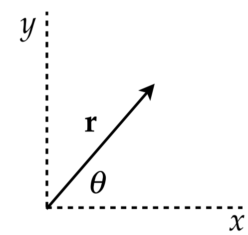
Any complex number z can be written in two forms, related by Euler's theorem:

$$z = x + iy$$
$$z = re^{i\theta}$$

In the first form, x and y are real numbers; x is the *real part* of z [written as $x = \mathrm{Re}(z)$] while y is the *imaginary part* of z [written as $y = \mathrm{Im}(z)$]. In the other form, r is the *amplitude* and θ is the *phase* of z. From Euler's theorem we find

$$x = r\cos\theta,\ y = r\sin\theta.$$

One can display a complex number graphically by showing the real and imaginary parts in a two-dimensional plot, as shown. Each point on the diagram corresponds to a particular complex number.



This two-component vector representing z is often called a *phasor* by engineers. Addition of two complex numbers is then accomplished by adding the phasors, using the usual rules for adding vectors.

Introductory physics textbooks often introduce phasors to treat superposition of oscillating functions and waves, without explaining that they are actually graphical representations of complex numbers. We will use the complex numbers directly.

The *complex conjugate* of a complex number is denoted by an asterisk, and is defined by replacing i by $-i$ wherever it appears. Thus

$$z^* = x - iy \ \text{ or } \ z^* = re^{-i\theta}$$

The product of z and z^* is a positive real number, and it is the square of the amplitude:

$$zz^* = x^2 + y^2$$
$$zz^* = r^2$$

Two more important facts about complex numbers:

- Two complex numbers are equal if and only if their real parts are equal *and* their imaginary parts are equal. This is two separate conditions.
- The real part of the sum of two complex numbers is the sum of their individual real parts. (The same is true of imaginary parts.)

Now we return to waves. We note that we could write the E-field of, say, the first wave in the form

$$E_1(x,t) = \mathrm{Re}[E_{0_1}e^{i(kx-\omega t)}]$$

and similarly for the other wave, and for the resulting wave.

Now we use the fact that the sum of real parts is the real part of the sum. We can temporarily replace the actual wavefunctions by the corresponding complex

exponential forms, add those, and then take the real part of the result to get the answer we want. That is the trick.

In our case, we define

$$E_1{}^c = E_{0_1} e^{i(kx-\omega t)}$$
$$E_2{}^c = E_{0_2} e^{i(kx-\omega t+\delta)}$$
$$E^c = E_0 e^{i(kx-\omega t+\phi)}$$

so that our actual wavefunctions are the real parts of these complex quantities. We add the first two and set the result equal to the third. Dividing out a common factor $e^{i(kx-\omega t)}$, we then find

$$E_0 e^{i\phi} = E_{0_1} + E_{0_2} e^{i\delta}$$

This equation between complex numbers is two equations, one for the real parts and one for the imaginary parts. These two equations allow us to solve for E_0 and δ.

The advantage of the Euler trick is that algebra of exponentials is much easier than algebra of sines and cosines.

To get the intensity of the resulting wave we need only $E_0{}^2$. We simply multiply each side of the above equation by its complex conjugate and use the fact that $e^{i\phi} \cdot e^{-i\phi} = 1$:

$$\begin{aligned} E_0{}^2 &= (E_{0_1} + E_{0_2} e^{i\delta})(E_{0_1} + E_{0_2} e^{-i\delta}) \\ &= E_{0_1}{}^2 + E_{0_2}{}^2 + E_{0_1} E_{0_2} (e^{i\delta} + e^{-i\delta}) \end{aligned}$$

From Euler's theorem we see that the ( ) in the last term is $2\cos\delta$, so we have

$$E_0{}^2 = E_{0_1}{}^2 + E_{0_2}{}^2 + 2E_{0_1} E_{0_2} \cos\delta .$$

This equation allows us to relate the intensity of the combined wave to the intensities of the original waves alone and the phase difference δ. We use the fact that for any e-m wave $I = KE_0{}^2$, where (if *I* represents the average intensity over a cycle) $K = \frac{1}{2} c\varepsilon_0$. We multiply every term of the above equation by *K*. Denoting the intensities of the original waves by I_1 and I_2, and calling the intensity of the combined wave *I*, we find:

Interference of two waves	$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos\delta$

This is the general formula for interference of two harmonic waves of the same wavelength and frequency, moving in the same direction but out of phase by δ.

The maximum intensity (constructive interference) occurs when δ is a multiple of 2π (i.e., when $\cos\delta = +1$). Minimum intensity (destructive interference) occurs when δ is an odd multiple of π (i.e., when $\cos\delta = -1$). We find from the above formula:

Constructive interference: $I = I_{\max} = \left[\sqrt{I_1} + \sqrt{I_2}\right]^2$

Destructive interference: $I = I_{\min} = \left[\sqrt{I_1} - \sqrt{I_2}\right]^2$

In many of the cases we will treat, the two waves have the same amplitude (therefore the same intensity alone), in which case the resulting wave has intensity

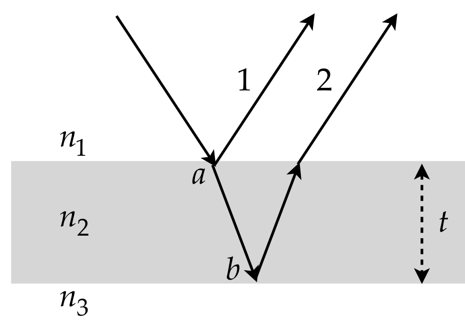
Waves of equal amplitude: $I = 2I_1(1+\cos\delta)$

This case occurs in most of our problems, so this is a very useful formula.

For this case, destructive interference gives *zero* intensity, while constructive interference gives *four* times the intensity of one wave alone.

The method outlined here will be generalized later to many waves, in the treatment of diffraction.

Interference in thin films



An important application of the above formulas is the case of light reflected from the two surfaces of a thin transparent film. Shown in cross section is an example, where the film is the shaded region.

Light is incident from above, coming from a medium with refractive index n_1. Wave 1 is the reflected wave from the top surface of the film. Part of the light enters the film, which has index n_2. Some of this is reflected at the bottom surface, and part of this emerges back into the original medium as wave 2. We are interested in interference between waves 1 and 2.

As was shown in the assignments, in the usual case where reflectivity R is small, these waves have approximately the same amplitude and other waves resulting from more reflections within the film have much smaller amplitudes than these two and can be neglected. We will assume this is the case.

The phase difference between these waves results from two causes:

- The difference in path followed by the two waves before they are brought together in a detector (e.g., the eye of an observer).
- Possible phase changes upon reflection at points *a* and *b* on the surfaces.

We assume the incident light is at nearly normal incidence. (The angles are exaggerated in the diagram for clarity.) In that case there are simple rules for any kind of wave, including light, about phase changes on reflection:

- If the waves have greater speed in the incident medium, then the reflected wave undergoes a phase change of π upon reflection.
- If the waves have lower speed in the incident medium, the reflected wave undergoes no phase change upon reflection.

In terms of light, if waves are incident *from* medium 1 *onto* medium 2:

Phase changes on reflection	1. If $n_1 < n_2$, the reflected wave changes phase by π. 2. If $n_2 > n_1$, there is no phase change in the reflected wave.

Now consider the case shown above. Waves 1 and 2 are out of phase because of the extra distance traveled by wave 2, and (perhaps) because of phase changes on reflections at *a* and *b*.

When light travels from one transparent medium to another with a different index of refraction, the frequency of the wave does not change. But because the wave speed changes, the wavelength $\lambda = v / f$ will be different, and so will the value of $k = 2\pi / \lambda$.

The ratio of *k* in the medium to that in the vacuum is

$$\frac{k}{k_{\mathrm{vac}}} = \frac{c}{v} = n$$

As wave 2 travels in medium 2, the phase of its wavefunction increases by $k_2 \cdot 2t = 4\pi n_2 t / \lambda$. This is the first reason for phase difference between waves 1 and 2.

We will always use λ to represent the *vacuum* wavelength.

Denote the phase change due to the reflection at *a* by δ_a, and that at *b* by δ_b. These numbers are either 0 or π, according to the rule given above. The net phase difference when the two waves come back together is thus

$$\delta = 4\pi n_2 t / \lambda + \left|\delta_b - \delta_a\right| .$$

This gives the value of δ to be substituted into the intensity formulas given earlier, to determine the reflected intensity.

Consider first the case where $n_1 < n_2$ and $n_2 > n_3$. (Example: a thin film with air on both sides.) Then $\delta_a = \pi$ and $\delta_b = 0$. This gives

$$\delta = 4\pi n_2 t / \lambda + \pi .$$

As the film thickness shrinks to zero (the film is about to break) $\delta \to \pi$, so there is *destructive* interference in the reflected light.

Now suppose that $n_1 < n_2$ and $n_2 < n_3$. (Example: A thin film of water on glass.) Then $\delta_a = \delta_b = \pi$ and the reflection phase changes cancel, giving

$$\delta = 4\pi n_2 t / \lambda .$$

As t goes to zero (the water evaporates) $\delta \to 0$ and there is constructive interference. If $n_2 t = \lambda / 4$, then $\delta = \pi$ and there is destructive interference.

In order for these effects to be observable the film must be no more than a few wavelengths thick. The reason for this is that no actual light source emits only a single wavelength; there is always a spread of wavelengths around the average. If the film is thick, so that the ratio t / λ is large, then a small change in λ will result in a large enough shift in δ to change destructive interference into constructive, or vice versa. These opposite patterns for nearby wavelengths overlap and blur each other's effects, so no overall pattern is visible. This is why we do not observe interference patterns in light reflected from the two surfaces of ordinary window glass.

This is analyzed in one of the assignments.